宇都宮大学教育学部　教育実践総合センター紀要　第29号　2006年7月1日

「ベンフォードの法則」の授業について†

井ノ口　順一*・黒﨑　貞雄**
宇都宮大学教育学部*
栃木県立宇都宮北高等学校**

高等学校の数学の授業をもっと魅力的なものに改善したいという考えから，この授業を企画した。授業の中で，自然現象，社会現象に現れる数値の最高位の数字に注目させ，どのような数字が出やすいのか調べさせた。その結果，１や２が出易いことが判明した。等比数列の例でこの法則（「ベンフォードの法則」）が成り立つ背景を数学的に説明できることを体験させた。

キーワード：数学教育，高等学校教育，ベンフォードの法則，最高位の数字，確率，対数

０．はじめに

「ベンフォードの法則」は，ある条件のもとでは数値データの先頭の数字として，１の出現頻度が１番高く，次いで２，３．．．の順に出やすいというものである。アメリカの物理学者Benford が定式化したといわれている［1］。（「対数分布定理」と呼ぶこともある。［4］）この法則は，会計処理の世界ではよく知られているようで，会計簿に現れる数値の先頭の数字の出現頻度に着目し，処理が適正かどうか，ねつ造されていないかをチェックするために活用しているとのことである［14］。かなり広い範囲で成り立つ法則のようなので［12］，自然現象や社会現象に現れる生のデータでチェックしてみるのも，魅力的な教材になるのではないかと考えた。いくつかの実践例も参考にさせていただいた［7］［8］［9］。数学的な内容はやや高度であり，改良の余地もあると思われるが，高校生のレベルで確かめられるということは意義のあることだと考える。

† Jun-ichi INOGUCHI*, Sadao KUROSAKI**
: Benford's law for high school students
*Department of Mathematics Education, Utsunomiya University
**Utsunomiya Kita High School

１．授業の概要

2006年2月7日宇都宮北高校３年理型の生徒を対象に，「ベンフォードの法則」を題材とした授業を行った。その概要を紹介する。

問題提起　Ｔ（教師）：『世の中の現象を表すデータに登場する数値の先頭の数字に注目しよう。一体どの数字が出やすいか。調べたり，考えたりしたことがありますか？』
Ｓ（生徒）：『ありません。』
Ｔ：『では，どの数字が一番多く出てくると思いますか。』
Ｓ：．．．．．。（見当が付かない様子。）

Ｔ：それでは，実際に調べてみよう。
まず，６種類の資料を用意した。

資料１：局部銀河群までの距離（理科年表2005年版より）

資料２：局部銀河群以外の明るい銀河までの距離（理科年表 2005 年版より）
資料３：単体の密度（理科年表 2005 年版より）
資料４：世界のおもな火山の標高（理科年表 2005 年版より）
資料５：ＯＥＣＤ諸国の 2003 年の国内総生産（単位：億ドル）（日本国勢図会 2005 年度版より）
資料６:2005 年の国連予算における分担率上位２４カ国（日本国勢図会 2005 年度版より）

　それぞれの資料のデータの先頭の数字に注目させた。
調べた結果は次のようであった。
資料１：最頻出数字は２，その頻度は
10/17＝0.588･･･
資料２：最頻出数字は１，その頻度は
13/36＝0.361･･･
資料３：最頻出数字は 1，その頻度は
38/114＝0.333･･･
資料４：最頻出数字は２，その頻度は
9/26＝0.346･･･
資料５：最頻出数字は 1，その頻度は
10/34＝0.294･･･
資料６：最頻出数字は１，その頻度は
9/24＝0.375･･･

Ｔ：『もう一つ，規則がわかっている例＝等比数列で調べてみよう。』
2^n（ｎ＝１，２，．．．，１００）の表を示し，先頭に出てくる数字の頻度を数えさせた。
結果は，表１のようになった。
Ｔ：『どうも１が出やすいようだ。数学を使って確かめよう。』

以下，生徒と一緒に議論を進めた。
　数列　$2^1, 2^2, 2^3, \ldots, 2^{100}$　について，先頭の数字が１となる確率を計算しよう。
　まず，先頭の数字が１となる条件を不等式で表すと，

$$100\cdots0 \le 2^k < 200\cdots0$$

指数で表せば

$$1\times10^m \le 2^k < 2\times10^m$$

（m は上式の０の個数）
という式が成り立つことである。k についての条件式を求めるために，各辺の常用対数をとると，

$$m \le k\log 2 < m + \log 2$$

ゆえに，先頭の数字が１となるために，k のみたすべき不等式は

$$\frac{m}{\log 2} \le k < \frac{m+\log 2}{\log 2}$$

よって，k が入る区間の幅を x_1 とすると，

$$x_1 = \frac{m+\log 2}{\log 2} - \frac{m}{\log 2} = 1$$

（なんと m の値によらない！）
同じようにして，先頭の数字が２となる条件は

$$2\times10^m \le 2^k < 3\times10^m$$

各辺の常用対数をとると，

$$m+\log 2 \le k\log 2 < m+\log 3$$

ゆえに，k の満たす不等式は

$$\frac{m+\log 2}{\log 2} \le k < \frac{m+\log 3}{\log 2}$$

よって，k が入る区間の幅は

$$x_2 = \frac{m+\log 3}{\log 2} - \frac{m+\log 2}{\log 2} = \frac{\log 3 - \log 2}{\log 2}$$

（これも m の値に関係ない！）
　先頭の数字が n となるために k がみたすべき（入るべき）区間の幅を

数字	1	2	3	4	5	6	7	8	9	計
頻度	0.3	0.17	0.13	0.1	0.07	0.07	0.06	0.05	0.05	1

表１：2^n（ｎ＝１，２，．．．，１００）の最高位の数字の分布

x_n ($n=1,2,.....,9$) とすると，上の議論と同様にして

$$x_3=\frac{m+\log 4}{\log 2}-\frac{m+\log 3}{\log 2}=\frac{\log 4-\log 3}{\log 2}$$

一般に

$$x_n=\frac{m+\log(n+1)}{\log 2}-\frac{m+\log n}{\log 2}=\frac{\log(n+1)-\log n}{\log 2}$$

$$(n=1,2,.....,9)$$

このとき

$$x_1+x_2+\cdots+x_9$$
$$=\frac{\log 2+(\log 3-\log 2)+(\log 4-\log 3)+\cdots(\log 10-\log 9)}{\log 2}$$
$$=\frac{1}{\log 2}$$

すると，先頭が１となる確率は

$$p_1=\frac{x_1}{x_1+x_2+\cdots+x_9}=\frac{1}{\left(\frac{1}{\log 2}\right)}=\log 2$$

先頭が n となる確率は

$$p_n$$
$$=\frac{x_n}{x_1+x_2+\cdots+x_9}$$
$$=\frac{\left\{\frac{\log(n+1)-\log n}{\log 2}\right\}}{\left(\frac{1}{\log 2}\right)}$$
$$=\log(n+1)-\log n=\log\left(1+\frac{1}{n}\right)$$

これを「ベンフォードの法則」という。

常用対数表をひくと，

先頭の数字が１となる確率＝$\log 2$＝0.3010

先頭の数字が２となる確率＝$\log 3-\log 2$

＝0.1761

先頭の数字が３となる確率＝$\log 4-\log 3$

＝0.1249

・・・・・・・・・

先頭の数字が９となる確率＝$\log 10-\log 9$

＝0.0458

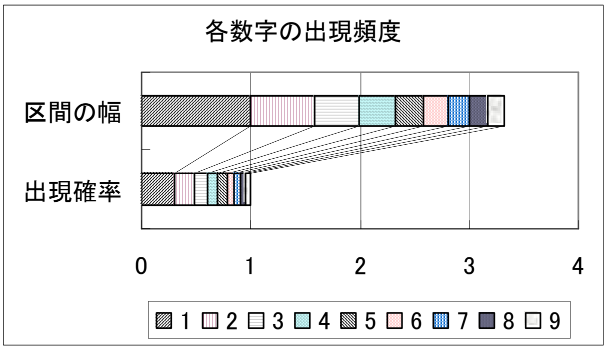

図１：ｋの満たす区間の幅と確率の関係

以上のような議論から，一番現れやすい数字は <u>１　</u>であり，その頻度は

<u>0.3010</u>である。

２．生徒の感想より

・Ｐ $_n$ ＝の計算がよくわからなかった。（後はよくわかった。）（ＲＳ）

- 計算が難しかった。１が出やすいのを初めて知った。
- よくわからなかったけど，授業は楽しかったです。
- いつもと違った「数学」に触れられて面白かったです。かなり難しかったけど．．．。たまにはこういったものをやると，数学のおもしろさがさらにＵＰすると思う！！１年間楽しかったです。☆ありがとうございました。！！（ＳＫ）
- 今日の授業は一言で言うとおもしろかったです。数学を学ぶ機会はこれから減るかもしれませんが，時間にゆとりがあるときにでも，積極的に数学に関わっていきたいと

思います。１年間ありがとうございました。

（ＴＷ）

- I don’t know. 分かりません。
- よくわかりませんでした。でもおもしろかったです。
- むずかしかった。わかんなかったけど，特別授業が最後の授業でおもしろかった。

３．検討

（１） 私立大学の入試本番を迎えつつある時期に，授業に積極的に参加した生徒がかなりいたことは収穫であった。この教材の魅力が十分にあることの証といえよう。

（２） 前半の問題提起の部分は，生徒のほとんどが興味を示してくれた。

（３） 一方，後半の数学的議論の部分は，難しく感じた生徒が多かった。この点が課題といえる。ｎの満たすべき区間の幅を利用して「確率」に持ち込むことに難しさを感じるようである。

（４） 上記の数学的議論の理解を助ける工夫を考えたい。一つには，視覚に訴えて図を利用すること。もう一つは，区間の幅が桁数に関係なく決まることを実感させること。

（５） どのような場合にこの「ベンフォードの法則」が成り立つのか，条件を確認する必要がある。参考資料［12］に挙げておいたホームページによると，ベキ乗で記述される法則(パワー則)が成り立つ場合には、スケールに関係なく普遍的であり、このときベンフォードの法則が成り立つとのことである。たとえば，万有引力，ほ乳類の酸素吸収量が体重の３／４乗に比例するという法則，個人の所得高の分布，都市人口を多い順に並べた分布，単語を出現頻度順に並べた分布などである。

（６） 2004 年度秋に宇都宮北高校学校２年生を対象に実践したときには，対象が文型の生徒であったため，興味を示してくれた生徒が多かった一方で，数学的議論の展開を十分に理解できた生徒が極めて少数であった。今回は理型の生徒を対象にしたため，困難さは大分違ったようである。

（７） 生データの資料として何を与えるか，あらかじめよく検討しておく必要がある。

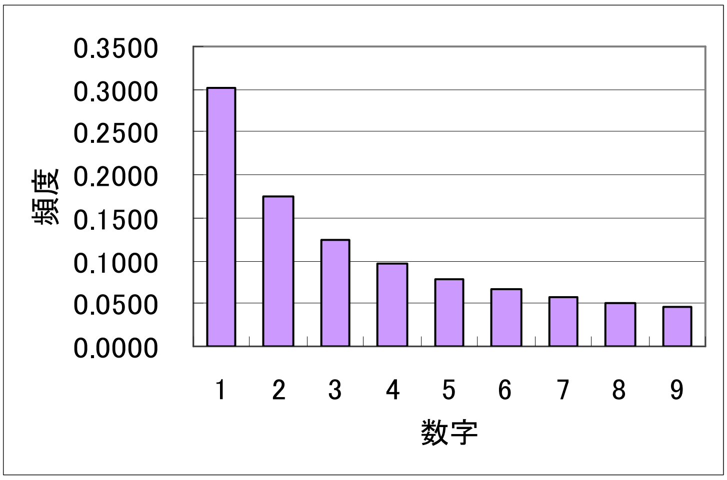

図２：各数字の出現頻度（確率）

４．補足

（１）ベンフォードの法則による理論値と比較したデータをあげてみる。

数字	理論値	2^n	フィボナッチ数列
1	0.301	0.300	0.300
2	0.176	0.170	0.180
3	0.125	0.130	0.130
4	0.097	0.100	0.090
5	0.079	0.070	0.080
6	0.067	0.070	0.060
7	0.058	0.060	0.050
8	0.051	0.050	0.070
9	0.046	0.050	0.040
計	1	1	1

表２：ベンフォードの法則による理論値

と各数列のはじめの 100 項における出現頻度

（２）数学的には次の定理が関係している［4]。

ヤコビの定理

『θ が無理数のとき，$\{m+n\theta \mid m,n\in Z\}$ は実数体 R において稠密である。』

一様分布定理（ワイル）

『θ が無理数のとき，$\{n\theta \mid n=1,2,\cdots\}$ の分布は一様である。

つまり、各 $x\in[0,1]$ に対し

$$\lim_{N\to\infty}\frac{\#\{n \mid 1\le n\le N,\{n\theta\}<x\}}{N}=x$$

が成り立つ。ここで $\{n\theta\}$ は $n\theta$ の小数部分を表している。』

５．参考文献および資料

［１］ F. Benford, The law of anomalous numbers,Proc.Amer.Phil.Soc.78(1938), 551-572.

［2］理科年表 2005 年版.

［3］日本国勢図会，2005 年版.

［4］吉田知行， １で始まる数が多いのはなぜか，数学の楽しみ No.１～４，日本評論社，1997.

［5］原田喜重編，数学っておもしろい，日本評論社，2001.

［6］根上生也，爽快！２¹⁰⁰三話，遊星社，1996.

［7］水野和明，統計学についての一考察，埼玉県立狭山清陵高等学校研究紀要.

［8］真鍋和弘， １で始まる数が多いのは本当か？，数学教育協議会全国研究札幌大会，2003.

［9］真鍋和弘， ７で始まる数　数論と確率の関係，数学教育協議会全国研究札幌大会，2003.

［10］井ノ口順一，ベンフォードの法則について，栃木数楽の会レポート，2003.

［11］黒﨑貞雄，“「ベンフォードの法則」１で始まる数値が多いのはなぜか？”，高等学校数学教育研究会紀要，2004.

［12］“パワー則とジップの法則”，htpp://www.geocities.jp/ikuro_kotaro/koramu/265 _zipf.htm

［13］“Benford’s Law”，http://mathworld.wolfram.com/BenfordsLaw.html

［14］“Following Benford's Law, or Looking Out for No. 1”，http://www.rexswain.com/Benford’sLaw.html

Summary

The purpose of this article is to exhibit an example (Benford’s law) of teaching materials in mathematics whichwould be attracitive for high school students.　In this lesson , students investigated the first digit of numbers in statistics on natural or social phenonema from high school mathematics viewpoints. As a result of this research , students found a mathematical fact that the number 1 and 2 tend to appear more frequently than other numbers . In case of geometric sequences , students confirmed the Benford’s law is corret by mathematical arguments.

Key Words：Mathematics education, High scholl education, Benford’s Law，First Digit, Probability，Logarithm.